3 Lecture 3

3.1 Separation

We are now going to move on to an extremely useful set of results called separation theorems. Basically, these theorems tell us that, if we have two convex sets, then we can stick a hyperplane between them. This should be intuitively obvious in $\mathbb{R}^2$. Remember that here, a hyperplane is simply a line, so the separation theorem simply tells us that we can draw a line between any two convex sets. However, this holds much more generally. For speed, we are going to work in $\mathbb{R}^n$, but you should be aware that these results hold in more general linear spaces.

Lets just begin with a reminder of what a hyperplane is in $\mathbb{R}^n$. We *defined* a hyperplane as a subset maximal affine manifold. However, in $\mathbb{R}^n$ we showed that any affine manifold could be written as the set

$$H(a, \alpha) = \{x \in \mathbb{R}^n | < x.a >= \alpha\}$$

for some $a \in \mathbb{R}^n/\emptyset$ and $\alpha \in \mathbb{R}$. The hyperplane is therefore defined by a, which we call the normal, and α.

We are going to start with a bunch of definitions of what it means to separate two sets

Definition 5 *let $X, Y \in \mathbb{R}^n$*

1. *X and Y are **separated** by $H(a, \alpha)$ if*

$$< x, a >\geq \alpha \geq < y, a > \ \forall\ x \in X,\ y \in Y$$

2. *X and Y are **properly separated** if they are separated and it is not the case that both $X \subset H(a, \alpha)$ and $Y \subset H(a, \alpha)$*

3. *X and Y are **strictly separated** by $H(a, \alpha)$ if*

$$< x, a >> \alpha >< y, a > \ \forall\ x \in X,\ y \in Y$$

4. *X and Y are **strongly separated** by $H(a, \alpha)$ if $\exists\ \varepsilon > 0$ such that $X + B(0, \varepsilon)$ and $Y + B(0, \varepsilon)$ are strictly separated*

It is worth noting a couple of other characterizations of strong separation. First, strong separation is the same as saying that for some ε

$$< x, a > -\varepsilon||a|| > \alpha >< y, a > +\varepsilon||a|| \; \forall \; x \in X, \; y \in Y$$

for some $\varepsilon > 0$. Alternatively, it must be the case that

$$\inf_{x \in X} < a, x >> \sup_{y \in Y} < a, y >$$

Example 1 *Some examples in $\mathbb{R}^2$ to fix ideas. Let $H = \left\{x \in \mathbb{R}^2 | x_2 = 0\right\}$*

1. *$X = [0, 2] \times \{0\}$ and $Y = [1, 3] \times \{0\}$ are separated by H but are not properly separated by H*
2. *$X = [0, 2] \times \{0\}$ and $Y = [1, 3] \times [0, 1]$ are properly separated by H but are not strictly separated by H*
3. *$X = [0, 2] \times [-1, 0)$ and $Y = [1, 3] \times (0, 1]$ are strictly separated by H but are not strongly separated by H*
4. *$X = [0, 2] \times [-1, -\varepsilon)$ and $Y = [1, 3] \times (\varepsilon, 1]$ are strongly separated by H*

More generally, note that

1. Sets can be separated from themselves (think of a line in $\mathbb{R}^2$)
2. A set cannot be properly separated from itself, but proper separation does not imply disjointedness
3. Two sets cannot be strictly separated unless they are disjoint

We can use the results that we proved previously to prove some separation results regarding convex sets. First, for a closed, convex set C, and y outside of C, we can find a hyperplane that strongly separates them.

Theorem 4 *Let $C \subset \mathbb{R}^n$ be convex, non-empty and closed, and $y \in \mathbb{R}^n / C$. Then there exists an $a \in \mathbb{R}^n / \emptyset$ such that*

$$< a, y >> \sup_{x \in C} < a, x >$$

(in other words y can be strongly separated from C by $H(a,\alpha)$ where $\alpha = \frac{1}{2} < a, y > + \frac{1}{2}\sup_{x\in C} < a, x >$)

Proof. *Let $x^* = P_c(y)$, and define $a = y - x^*$. We know that*

$$< y - x^*, x - x^* > \leq 0 \;\forall\; x \in C$$

or

$$\begin{aligned} < & \; a, x + a - y > \leq 0 \\ ||a||^2 + < & \; x, a > - < a, y > \leq 0 \\ < & \; a, y > \geq ||a||^2 + < x, a > \end{aligned}$$

Which, as $a \neq \emptyset$ implies $< a, y >> \sup_{x\in C} < a, x >$ ■

In fact, the result is stronger this: We can also strongly separate *sets*

Theorem 5 *Let $C_1, C_2 \subset \mathbb{R}^n$ be convex. Assume that C_1 is closed and C_2 is compact and the two sets are disjoint. Then they can be strongly separated.*

Proof. *Let $K = C_1 - C_2$. You will show for homework that this is a closed and convex set. Also, as $C_1 \cap C_2 = \emptyset$, $\emptyset \notin K$.*

There therefore exists (by theorem 4) some $a \in \mathbb{R}^n$ such that $a \neq 0$ and

$$\begin{aligned} 0 &= < a, 0 >> \sup_{k\in K} < a, k >= \sup_{\substack{c_1\in C_1 \\ c_2\in C_2}} < a, c_1 - c_2 > \\ &= \sup_{c_1\in C_1} < a, c_1 > - \inf_{c_2\in C_2} < a, c_2 > \\ &\Rightarrow \inf_{c_2\in C_2} < a, c_2 >> \sup_{c_1\in C_1} < a, c_1 > \end{aligned}$$

■

A natural question is whether we need all those restrictions. Clearly the result is not if and only if: there are non-closed non-compact sets that we can strongly separate. But can we *guarantee* strong separation without all these restrictions? The answer is no. In particular, if we drop the condition that C_2 has to be bounded, we lose strong separation, as you will show for homework.

So what do we get if we weaken these conditions - can we still say anything? It turns out that the answer is yes. Returning to the case of separating a point from a set, it turns out that we can still get proper separation while only assuming convexity

Theorem 6 *Let $C \subset \mathbb{R}^n$ be convex and $y \in \mathbb{R}^n/C$. Then there exists a hyperplane $H(a, \alpha)$ that properly separates y from C. i.e.*

$$< a, y >= \alpha \geq < a, x > \ \ \forall \ x \in C$$

Moreover, $\alpha >< a, x > \forall \ x \in int(C)$

Proof. *Beyond scope of this course* ■

A corollary to this theorem is that, for any open, convex set C and any point outside of the set, there is a hyperplane such that $< a, y >= \alpha$ and $\alpha >< a, x > \forall \ x \in C$

Corollary 3 *For any open, convex set C and any point outside of the set, there is a hyperplane such that $< a, y >= \alpha$ and $\alpha >< a, x > \forall \ x \in C$*

Before we move on, there is one final result that will prove useful

Theorem 7 *Let $A \subset \mathbb{R}^n$ be contained in some half-space. Then $cl(co(A))$ is the intersection of all closed half spaces that contain A*

Proof. *Let $H_-(a, \alpha) = \{x \in \mathbb{R}^n | < a, x > \leq \alpha\}$, and let*

$$\Omega = \{(a, \alpha) \, | a \neq 0 \ \ and \ A \subset H_-a, \alpha\}$$

If $A \subset H_-(a, \alpha)$, then as $H_-(a, \alpha)$ is convex, $co(A) \subset H_-(a, \alpha)$, and so $cl(co(A)) \subset cl(H_-(a, \alpha)) = H_-(a, \alpha)$

Now take a point $x \neq cl(co(A))$. Then there exists (a, α) such that

$$< a, x > \alpha >< a, z > \ \ \forall \ z \in cl(co(A))$$

So $A \subset H_-(a,\alpha)$ *and* $x \notin H_-(a,\alpha)$, *so*

$$x \notin \cap_{(a,\alpha)\in\Omega} H_-(a,\alpha)$$

■